

REMARKS

Claims 1-6 and 8 are pending in this application after this amendment. Claims 1, 6, and 8 are independent. Claim 7 has been canceled without prejudice or disclaimer to the subject matter included therein. In light of the amendments and remarks made herein, Applicant respectfully requests reconsideration and withdrawal of the outstanding rejections.

By this amendment, Applicant has amended the claims to more appropriately recite the present invention. These amendments are being made without conceding the propriety of the Examiner's rejections, but merely to timely advance prosecution of the present application.

In the outstanding Official Action, the Examiner rejected claim 7 under 35 U.S.C. §101; rejected claims 1 and 4-8 under 35 U.S.C. §102(e) as being anticipated by Stevens et al. (U.S. Patent Application Publication No. 2002/0138265); and rejected claims 2 and 3 under 35 U.S.C. §103(a) as being unpatentable over Stevens et al. in view of Chen et al. (USP 6,006,186). Applicant respectfully traverses these rejections.

Claim Rejections - 35 U.S.C. §101

The Examiner rejected claim 7 asserting the claim is directed to non-statutory subject matter. By this amendment, Applicant has canceled claim 7. As such, it is respectfully requested that the outstanding rejection be withdrawn.

Rejection under 35 U.S.C. §102

Claims 1 and 4-8 stand rejected under 35 U.S.C. §102(b) as being anticipated by Stevens et al. In support of the Examiner's rejection of claim 1, the Examiner asserts that Stevens et al. discloses a context dependent acoustic model storage unit storing context dependent acoustic models in a form of sub-word state trees. In support of this assertion, the Examiner asserts that Stevens et al. discloses "each phoneme may be represented as a triphone that includes multiple nodes. A triphone is a context-dependent phoneme." The Examiner additionally cites to

paragraph [0075]. Applicant respectfully disagrees with the Examiner's characterization of this reference.

The disclosure of Stevens et al. is directed to error correction in speech recognition. In paragraphs [0075] – [0078], Stevens et al. discloses as follows:

[0075] The active vocabulary 230 uses a pronunciation model in which each word is represented by a series of phonemes that comprise the phonetic spelling of the word. Each phoneme may be represented as a triphone that includes multiple nodes. A triphone is a context-dependent phoneme. For example, the triphone "abc" represents the phoneme "b" in the context of the phonemes "a" and "c", with the phoneme "b" being preceded by the phoneme "a" and followed by the phoneme "c".

[0076] One or more vocabulary files may be associated with each user. The vocabulary files contain all of the words, pronunciations, and language model information for the user. Dictation and command grammars may be split between vocabulary files to optimize language model information and memory use, and to keep each single vocabulary file under 64,000 words.

[0077] Separate acoustic models 235 are provided for each user of the system. Initially speaker-independent acoustic models of male or female speech are adapted to a particular user's speech using an enrollment program. The acoustic models may be further adapted as the system is used. The acoustic models are maintained in a file separate from the active vocabulary 230.

[0078] The acoustic models 235 represent phonemes. In the case of triphones, the acoustic models 235 represent each triphone node as a mixture of Gaussian probability density functions ("PDFs")...

While Stevens et al. discloses a triphone that includes multiple nodes, Stevens et al. discloses the acoustic models representing each triphone node as a mixture of Gaussian probability density functions. There is no disclosure that is directed to a context dependent acoustic model storage unit in which the context dependent acoustic models **are stored in a form of sub-word state trees** in each of which state sequences of a plurality of sub-word models of the context dependent acoustic models are organized in a tree structure, as required by the claim.

Further, the Examiner asserts that Stevens et al. discloses a matching unit developing hypotheses of sub-words by referencing the sub-word state tree representing the context dependent acoustic models, the word lexicon and the language models, **and performing matching between feature parameters of inputted speech and the developed hypotheses so as to output word information including a word, an accumulated score and a beginning start frame with respect to a hypothesis representing a word end portion.** In support of this assertion, the Examiner relies on paragraphs [0060] and [0169]. Applicant respectfully disagrees with the Examiner's characterization of this reference.

Stevens et al. discloses in paragraph [0060] as follows:

[0060] A recognizer 215 receives and processes the frames of an utterance to identify text corresponding to the utterance. The recognizer entertains several hypotheses about the text and associates a score with each hypothesis. The score reflects the probability that a hypothesis corresponds to the user's speech. For ease of processing, scores are maintained as negative logarithmic values. Accordingly, a lower score indicates a better match (a high probability) while a higher score indicates a less likely match (a lower probability), with the likelihood of the match decreasing as the score increases. After processing the utterance, the recognizer provides the best-scoring hypotheses to the control/interface module 220 as a list of recognition candidates, where each recognition candidate corresponds to a hypothesis and has an associated score. Some recognition candidates may correspond to text while other recognition candidates correspond to commands. Commands may include words, phrases, or sentences.

Further, Stevens et al. discloses in paragraph [0169] as follows:

[0169] Scores for confused pronunciation matches in the general phoneme confusability matrix may be generated using three sources of information: the probability that a sequence of phonemes for which the matches were sought (a recognized sequence) was the actual sequence of phonemes produced by the speaker, the probability that a particular confused pronunciation (the confused sequence) was confused for the recognized sequence, and the probability that the confused sequence occurs in the language (for example, English) with which the speech recognition system is used. These probabilities correspond to the scores produced by, respectively, the recognizer for the recognized sequence, a dynamic programming match of the recognized phonemes with the dictionary pronunciation using a priori probabilities of phoneme confusion, and an examination of a unigram language model for the words corresponding to the pronunciation of the recognized sequence.

As can be seen from the above disclosure, Stevens et al. teaches a recognizer 215 that generates a best-scoring hypotheses to the control/interface module 220 as a list of recognition candidates, where each recognition candidates corresponds to a hypothesis and has an associated score. However, Applicant maintains that Stevens et al. fails to disclose performing matching between feature parameters of inputted speech and the developed hypotheses so as to output word information including a word, an accumulated score and a beginning start frame with respect to a hypothesis representing a word end portion, as required by the claim.

As Stevens et al. fails to teach or suggest all of the claim elements, Applicant maintains that claim 1 is not anticipated by, and thus allowable over, the teachings of Stevens et al. It is respectfully requested that the outstanding rejections be withdrawn.

Claims 2-5 are allowable for the reasons set forth above with regard to claim 1 at least based on their dependency on claim 1. Further, claims 6 and 8 include elements similar to those discussed above with regard to claim 1 and thus these claims are allowable for the reasons set forth above with regard to claim 1.

Conclusion

In view of the above remarks, it is believed that claims are allowable.

Should there be any outstanding matters that need to be resolved in the present application, the Examiner is respectfully requested to contact Catherine M. Voisin Reg. No. 52,327 at the telephone number of the undersigned below, to conduct an interview in an effort to expedite prosecution in connection with the present application.

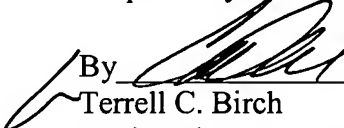

Application No. 10/501,502
Amendment dated November 15, 2007
Reply to Office Action of August 15, 2007

Docket No.: 0020-5278PUS1

If necessary, the Commissioner is hereby authorized in this, concurrent, and future replies to charge payment or credit any overpayment to Deposit Account No. 02-2448 for any additional fees required under 37.C.F.R. §§1.16 or 1.14; particularly, extension of time fees.

Dated: November 15, 2007

Respectfully submitted,

By  
Terrell C. Birch
Registration No.: 19,382 #52327
BIRCH, STEWART, KOLASCH & BIRCH, LLP
8110 Gatehouse Road
Suite 100 East
P.O. Box 747
Falls Church, Virginia 22040-0747
(703) 205-8000
Attorney for Applicant